TM 79-3 SY

Technical Memorandum

VOICE STRESS ANALYSIS AS A
MEASURE OF OPERATOR WORKLOAD

S. G. Schiflett, Ph.D
G. J. Loikith, B. S.

Systems Engineering Test Directorate

10 March 1980

Reproduced From
Best Available Copy

20010531 100

Approved for Public Release; Distribution Unlimited.

NAVAL AIR TEST CENTER
PATUXENT RIVER, MARYLAND

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TM 79-3 SY	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* VOICE STRESS ANALYSIS AS A MEASURE OF OPERATOR WORKLOAD		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL MEMORANDUM FEB 1978 - DEC 1979 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) S. G. SCHIFLETT, Ph.D G. J. LOIKITH, B. S.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS NAVAL AIR TEST CENTER NAVAL AIR STATION PATUXENT RIVER, MARYLAND 20670		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS NAVAL AIR SYSTEMS COMMAND DEPARTMENT OF THE NAVY WASHINGTON, D.C. 20361		12. REPORT DATE 10 MARCH 1980 13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* NAVAL AIR TEST CENTER SYSTEMS ENGINEERING TEST DIRECTORATE (SY721) PATUXENT RIVER, MARYLAND 20670		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

WORKLOAD
STRESS
VOICE
OPERATOR
PERFORMANCE
MEASUREMENT
TECHNIQUE
ANALYSIS
INFORMATION PROCESSING

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This study attempted to determine if the Psychological Stress Evaluator (PSE) could be used to detect the amount of situational stress in the voice while subjects performed a four-choice information processing task at different presentation rates. The 42 subjects were divided into group I - Jet, group II - Prop, and group III - Staff. A Response Analysis Tester (RATER) presented a four-choice discrimination task in which the subjects were required to match a response key to each of four stimuli (numbers - one, two, three, and four) appearing in a display window. The sequence of stimuli was randomly presented in an automatic-paced mode for nine 1-min tests. The stimuli presentation rates were set

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

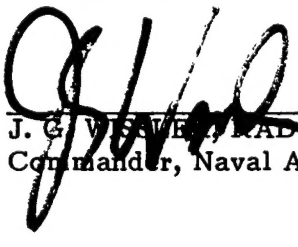
20.

at one symbol per 1.5 sec, .75 sec, and .50 sec. During the first three tests, the subjects were instructed to press the correct key and not verbalize the number. During the next three tests, the subjects verbalized the number and did not press the key. During the last three tests, the subjects verbalized the number and simultaneously pressed the correct key. At the end of each block of three tests, the subjects estimated self-performance as percent correct and rated stress on a scale of one (no stress) to seven (high stress). Voice signals were initially recorded on magnetic tape, then processed through filtering circuits and displayed on a strip chart for subsequent visual analysis and interpretation. A subjective scoring criterion was established and then translated into electronic equivalents and automated on a Varian 73 computer for voice pattern recognition analysis. Significant main effects for percent-correct responses were obtained for groups, presentation rate, and groups X presentation rate interaction. No significant differences were found in the correct responses of the subjects when the number was verbalized or not verbalized. The Staff group produced significantly fewer correct responses than either the Jet or Prop groups at the .75 sec rate. Voice stress analysis showed significant correlations with performance scores and stress ratings of a selected pool of subjects (N=12). The results were discussed as to the potential application of an objective, reliable, sensitive, and nonobtrusive measure of stress in vocal communication systems that require operator workload assessments.

PREFACE

This technical memorandum documents a study conducted at the Naval Air Test Center from February 1978 to December 1979. The technical investigation was conducted as a continuing effort of the Aircrew Systems Branch to support development of Human Factors test and evaluation methodology. Specifically, the project was initiated because of a recommendation to quantitatively determine operator workload levels in high stress environments (reference NAVAIRTESTCEN RW-9R-78).

APPROVED FOR RELEASE



J. G. [illegible], RADM, USN
Commander, Naval Air Test Center

TABLE OF CONTENTS

	Page No.
REPORT DOCUMENTATION PAGE	i
PREFACE	iii
TABLE OF CONTENTS	iv
INTRODUCTION	1
NEED	1
BACKGROUND	1
PURPOSE	3
METHOD	4
SUBJECTS	4
DESIGN	4
TASK/EQUIPMENT	4
SCORING STRATEGY	5
RESULTS	7
PERFORMANCE SCORES (% CORRECT RESPONSES)	7
SELF-RATED STRESS	8
VOCAL STRESS ANALYSIS	9
CORRELATIONS	10
DISCUSSION	12
CONCLUSIONS	14
RECOMMENDATIONS	15
APPENDIX A - AUTOMATED VOCAL STRESS SCORING TECHNIQUE	16
REFERENCES	21
DISTRIBUTION	23

INTRODUCTION

NEED

1. A need exists for a general stress measure which is nonobtrusive and can be applied in actual in-flight operational situations without interfering with operator performance. Ideally, the technique should be objective and quantifiable, simple to collect and analyze, reliable across individuals, and valid for specific aircrew operational test situations.

BACKGROUND

2. In several recent surveys of potential operator workload techniques (Gartner & Murphy, 1976; Harrison, 1977; Roscoe, 1978; and Wierwille, Williges & Schiflett, 1979), it was found that no single technique could be recommended as the definitive measure of operator workload because of the multidimensionality of stress inherent in an operational setting. An additional conclusion reached by Schiflett (1976) was that the majority of techniques developed were used in the design stage of aircrew systems, thereby making them difficult and/or impractical to implement in the later stages of operational test and evaluation.

3. However, it is very apparent from observational data that performance often involves some form of verbal communication associated with completion of a task. Speech may be the common denominator underlying the search for a general stress measure which is nonintrusive in the operational testing environment. If it can be demonstrated that stress is manifested in the acoustic speech signal, it may be possible to develop an automatic measurement technique for analyzing a speaker's vocalizations to determine when and to what degree stress is present.

4. Previous research (Lieberman, 1961; Lieberman & Michaels, 1962; and Bell, Ford & McQuiston, 1972) in the field of voice analysis has indicated that the psychophysiological state-of-the-speaker may manifest itself in the acoustic domain by changes in the frequency modulation (8 to 12 Hz) of the fundamental frequency (50 to 250 Hz) of the voice. An electronic instrument called the Psychological Stress Evaluator (PSE) was used to detect, analyze, and display these frequency changes in the inaudible microtremors of the muscles that surround the vocal cords. The damping of these muscular undulations is reported to be associated with mild tension by Lippold (1971), Heisse (1974), and Inbar (1976) even though there is some question to the origin of the physiological mircotremors (McGlone, 1975). However, in a thorough review of the research associated with speech patterns in aircrew workload, Cannings (1979) recommended that the use of microtremor would be a useful approach and warrants further exploratory investigations.

5. The majority of validation studies reviewed by Kubis (1974), Edson (1976), Heisse (1976), and Hirsch & Wiegele (1979) was conducted in clinical settings, laboratory deception-detection experiments, or actual criminal case histories involving lie detection of known solutions either by confession or court decisions. The results, with the exception of the Kubis report, generally support the validity

of the instrument to detect stress associated with deception. However, several serious methodological problems are discussed by Brenner & Branscomb (1979) which directly affect its use in lie detection settings.

6. The PSE instrument has also been used in a limited quantity by a wide variety of applications other than lie detection. Several annotated bibliographies are available from Dektor, Incorporated, that provide the reader with a good review of nondeception research studies ranging from clinical applications to political speech analysis. An attempt will not be made to review all of these nondeception studies but only those that use a research paradigm or actual field study to induce task-related stress in individuals.

7. A series of preliminary studies was conducted by Planar Corporation (1972) to test the feasibility of using the PSE to measure stress attributable to workload imposed by task demands and/or perceived risk. Even though the sample sizes were small and the task situations were not under experimental control, the findings did suggest that the PSE could be used to quantify vocal stress in aircraft accident investigations, air traffic controller messages, and astronaut communications during space missions. An in-depth, follow-up study by Older & Jenney (1975) analyzed voice communications of Skylab astronauts engaged in operational tasks of varying degrees of difficulty, e.g., conduct of solar observations and study of earth resources. Although some statistically significant relationships were found, the voice analysis technique was not judged to be sufficiently predictive to warrant its use in assessing the degree of psychological stress of crew members in future space missions.

8. Favorable validation evidence from a battery of psychophysiological measures has been reported by Borgen & Goodman (1976) of the Parke-Davis laboratories. The results yielded systematic changes in PSE scores to the stroop color/word conflict task. The PSE scores also correlated well with changes in heart rate, blood pressure, skin potential, and forearm blood flow.

9. Brenner, Branscomb & Schwartz (1979) tested the PSE on a mental arithmetic task which had previously produced graded changes in heart rate, pupil dilation, and skin resistance (Kahnenan, Tursky, Shapiro, & Crider, 1969). The subjects performed a series of mental arithmetic problems which varied in difficulty but which had a fixed execution time. The PSE scores increased directly with task difficulty and paralleled error data, self-report data, and previous psychophysiological evidence. Although shortcomings of the PSE scores were noted, especially subjectivity in scoring and response-word artifacts, the overall results suggest the basis for a practical measure of stress.

10. Encouraged by these results, Branscomb (1979) and Brenner, Branscomb, & Wright (1979) developed a computer-based measure of vocal stress by automating the PSE scoring procedures and correlating the results with pupil diameter changes while performing a mental arithmetic task. A multiple regression coefficient of $r = .39$ was obtained by combining all nine vocal measures to predict task workload. Coefficients for individual subjects were higher, ranging in prediction from 11% to 45% of the available variance. The researchers concluded that, given the preliminary nature of the present vocal measure algorithms, the findings are clearly encouraging and warrant further investigation.

PURPOSE

11. The paper documents a study presented by Schiflett (1979) that attempted to determine if the PSE could be used to detect the amount of situational stress in the voice while subjects performed a four-choice information processing task at different presentation rates. The purpose was to evaluate the relationship between performance scores, self-rated stress, and measured vocal stress due to increases in information presentation rates.

METHOD

SUBJECTS

12. The subject pool consisted of 31 male Naval officers and 11 male civilians, all college graduates between the ages of 24 and 39. At the time of testing, the participants were all enrolled in a graduate degree program sponsored by the Naval Air Test Center, Patuxent River, Maryland.

DESIGN

13. The effects of presentation rate on performance and stress were investigated using a 3 X 3 X 2 Analysis of Variance (ANOVA) factoral design (Kirk, 1968) and nonparametric statistics (Siegel, 1956). The military officers were divided into two groups based on the type of flight squadron experience (i.e., attack/fighter or multi-engine/helicopter). The third group was composed of military officers in staff (nonflight) billets and civilian personnel. The groups were designated as group I - Jet, group II - Prop, and group III - Staff. Subjects were randomly assigned, within each group, to the order of tests.

14. A subgroup of 12 subjects was selected for vocal stress analysis out of the larger (N=42) pool of subjects. The criterion for selection was the rank order of percent-correct scores, i.e., four highest, four lowest, and four middle-ranked scores averaged across all presentation rates. These subjects were selected irrespective of group identity.

TASK/EQUIPMENT

15. The psychomotor tests were displayed to the subject and automatically scored on a Response Analysis Tester (RATER). The RATER presented a four-choice discrimination task in which the subject was required to match a response key to each of four stimuli (numbers - one, two, three, and four) appearing in a display window.

16. The sequence of stimuli was randomly presented in an automatic-paced mode for nine 1-min tests. Each subject received three different presentation rates under three response conditions. The stimuli presentation rates were one symbol per 1.5 sec, .75 sec, and .50 sec for 1 min each. During Condition 1 (key presses only), the subjects were instructed to press the correct key and not vocalize the number. During Condition 2 (vocal only), the subjects vocalized the number and did not press the key. During Condition 3 (key presses and vocalizations), the subjects vocalized the number and simultaneously pressed the correct key. All conditions were randomized to control order effects. At the end of three tests of each condition, the subjects estimated self-performance as percent correct and rated stress on a scale of one (no stress) to seven (high stress) for each presentation rate. The subjects were allowed as many practice sessions as needed to reach a criterion of 90% correct response at the 1.5-sec rate for Condition 3 (key pressed and vocalization). All subjects reached the criterion within three 1-min sessions. After practice, a 1-min pretest was given using a self-paced mode feature that allowed the subject to control the presentation rate of the stimuli by responding as quickly

and accurately as possible. The pretest was used as a measure of each individual's baseline responding rate. Also, a self-pacing posttest was given at the end of the nine 1-min tests to determine any differences in mean deviations from pretest baseline conditions.

17. The voice was analyzed by a PSE manufactured by Dektor, Incorporated, developed specifically as a deception-detection instrument. The device consists of a signal analyzer, a strip chart pen recorder, a magnetic tape recorder, and a microphone. Voice signals were initially recorded on magnetic tape, then processed through the analyzer circuits, and displayed on a strip chart for subsequent visual analysis and interpretation.

SCORING STRATEGY

18. Interpretation of the strip chart tracing is typically accomplished by visual examination of the recorded signal for specific patterns. Comprehensive training is necessary to identify the most important characteristics of each pattern in order to relate the output signals to a level of stress. All PSE voice charts were coded by group number, presentation rate, and response condition, i.e., with or without key presses. A scoring criterion of waveform uniformity was developed from the coded charts by two interpreters trained to recognize stress patterns and response word artifacts in the PSE charts. The scoring criterion of uniformity was quantified by simulating vocal stress patterns as shown in figure 1a. These patterns were generated by an analogue waveform synthesizer and fed into a Varian 73 computer via an analogue-to-digital converter. The peak-to-peak displacement scoring algorithm was detected by sampling at 2,000 times-per-second. The sample waveforms were scored and ranked for degree of stress/uniformity that yielded a scale from 0 (low stress/uniformity) to 10 (high stress/uniformity). The uniformity scale was then used to quantify the actual measured vocal stress patterns for each selected test subject.

19. Representative samples of actual vocal patterns indicating a range of stress are shown in figure 1b. Notice the degree of change in the uniformity, i.e., peak-to-peak differences of the sample waveforms. The lack of uniformity or more randomization in the pattern indicates less stress manifested in the vocal output of the subject. These uniformity changes are quantified in units of stress as shown in the sample patterns.

20. The automated scoring technique, as shown in more detail in appendix A, was used to measure the degree of stress in the waveform pattern identified by the PSE from the vocal utterances of each selected test subject. The measurement technique provided a reliable and expedient method of scoring large quantities of data. For example, each of the selected 12 subjects, for nine 1-min tests and the pre-posttests yielded approximately 850 vocal utterances to be analyzed. Since the stimulus numbers were presented randomly within each presentation rate, a sampling method was devised to eliminate approximately one-half of the data by selecting every other utterance at the .75 sec/symbol rate and every third utterance at the .50 sec/symbol rate. All vocalizations at the slow rate of 1.5 sec/symbol and the pre-post conditions were analyzed without sampling.

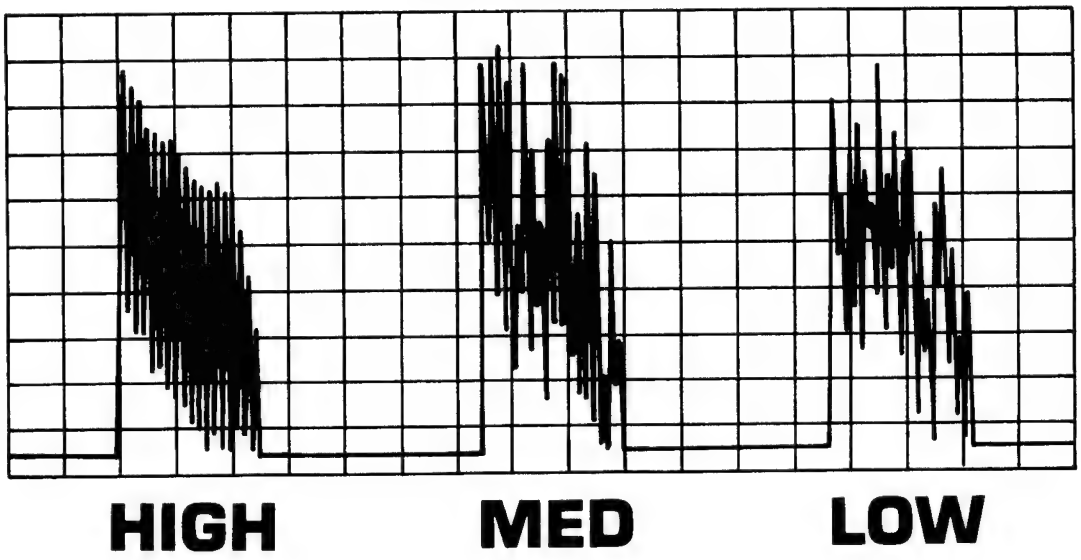


Figure 1a
Simulated Vocal Stress Patterns

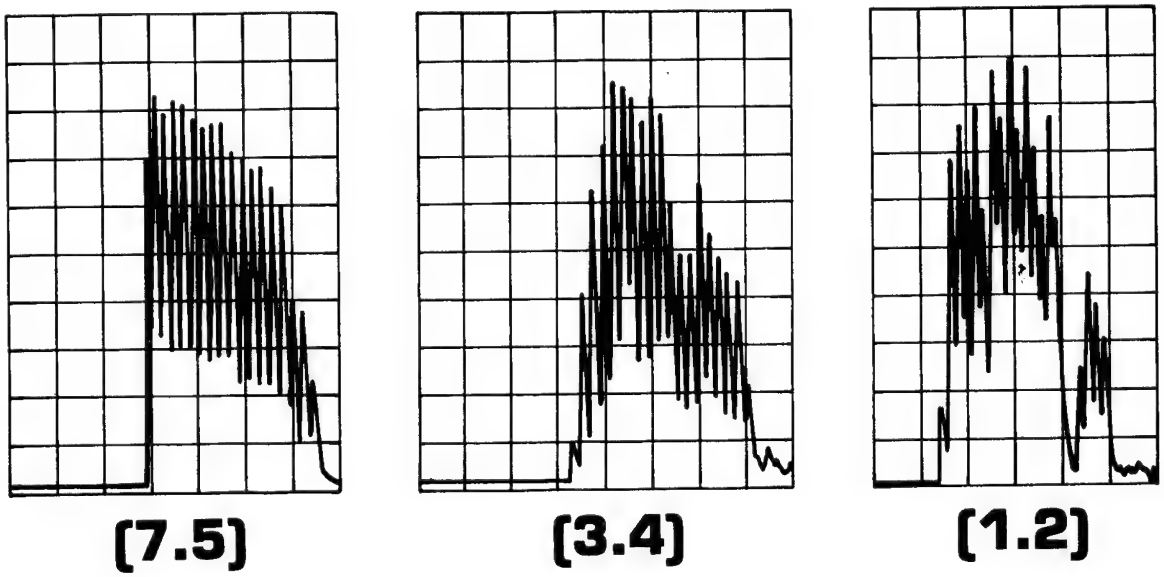


Figure 1b
Measured Vocal Stress Units

RESULTS

PERFORMANCE SCORES (% CORRECT RESPONSES)

21. The data for percent-correct responses for all subjects under each experimental condition are presented in figure 2. Significant main effects were obtained from ANOVA for groups ($F = 6.02$, $df = 2/234$, $p < .05$) and presentation rate $F = 1308.27$, $df = 4/234$, $p < .05$). No significant differences were found in the correct responses of the subjects when the number was vocalized or not vocalized ($F = 0.41$, $df = 1/234$, $p > .05$).

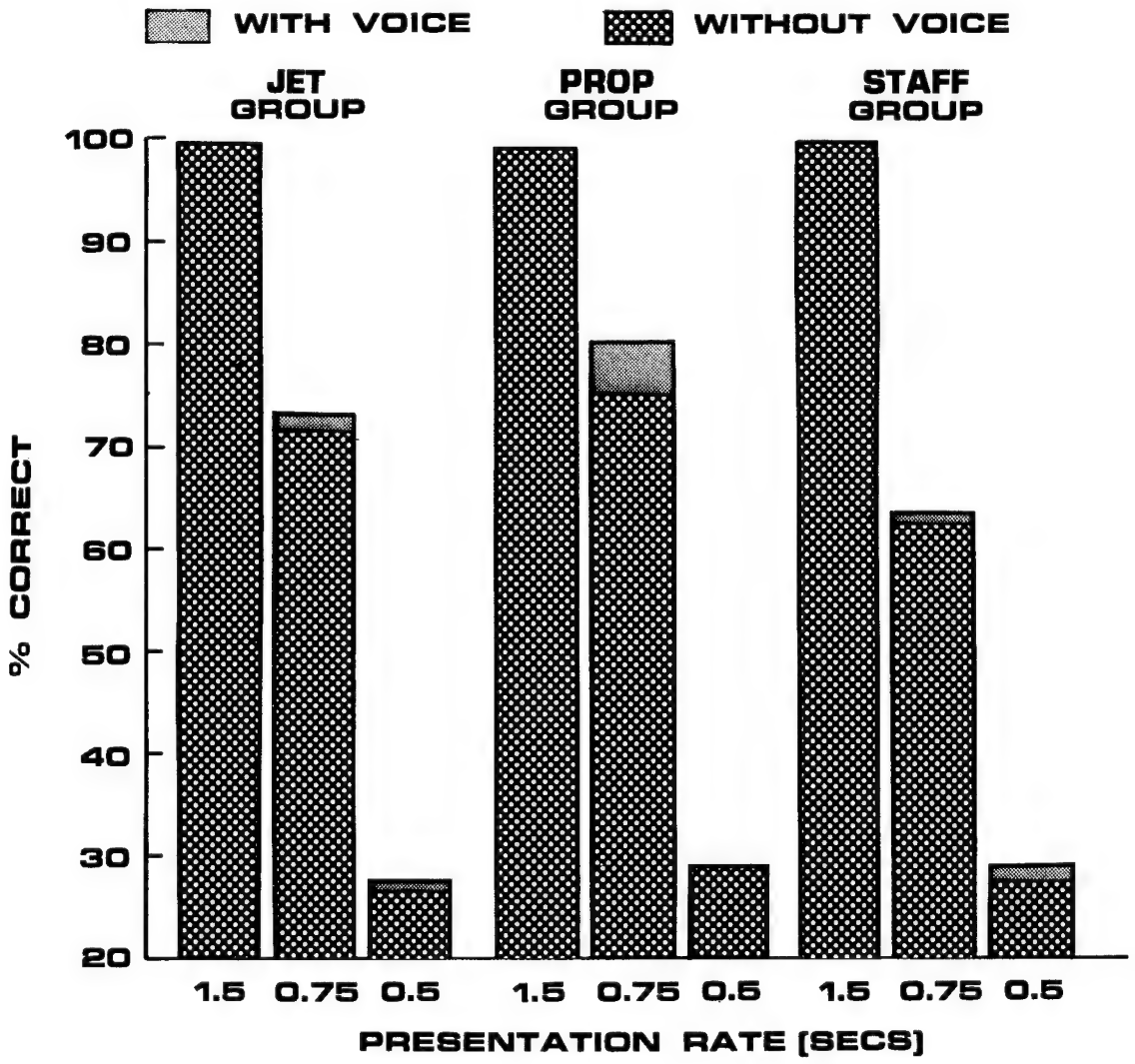


Figure 2
Percent Correct Key Presses

22. Tests of simple main effects produced significant differences across groups at each presentation rate. A series of Tukey's mean comparison tests revealed that presentation rate means were all significantly different from each other within each group.

23. No differences were obtained between groups except at the .75-sec presentation rate. The Staff group (X = 62.93%) produced significantly lower correct responses than the Jet (X = 72.31%) or Prop (X = 77.40%) groups. There were no significant differences in percent correct scores between the Jet and Prop groups.

24. The percent correct estimates by the subjects did not show any significant differences across groups or response conditions. The estimates of performance were significantly affected by presentation rate. Performance was consistently estimated to be lower than actual percent-correct scores at the 1.5-sec and .75-sec rates. The subjects significantly gave higher estimates of performance at the .5-sec rate than actual measured percent-correct scores.

SELF-RATED STRESS

25. The data for self-rated stress for all subjects under each experimental condition are presented in figure 3. Significant main effect differences were obtained for groups, presentation rates, and response condition. No significant interactions were found between any of the main effect variables on the self-rated stress measure.

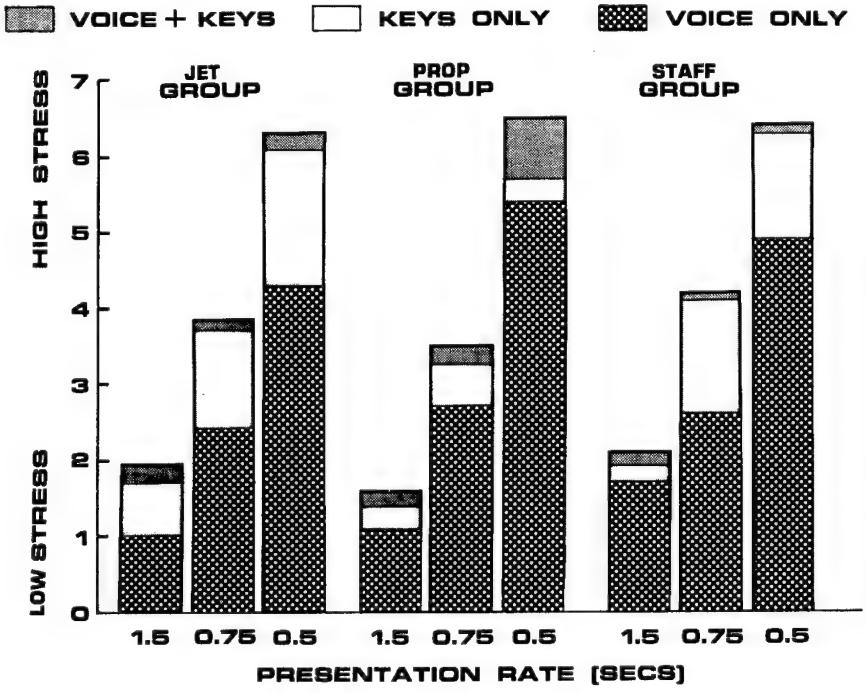


Figure 3
Self-Rated Stress

26. The Staff group ($\bar{X}$ = 3.72) rated stress slightly but significantly (F = 4.03, df = 2/351, p < .05) higher than either the Jet group ($\bar{X}$ = 3.41) or Prop group ($\bar{X}$ = 3.46) while performing under all response conditions.

27. Presentation rate significantly (F = 673.77, df = 2/351, p < .05) affected self-ratings for all groups whether vocalizing the numbers only, key pressing only, or in combination. The self-rated stress means for all groups at each presentation rate were: (a) low rate (1.49), (b) medium rate (3.32), and (c) high rate (5.79).

28. All groups significantly (F = 54.36, df = 2/351, p < .05) rated the amount of stress in the task of responding vocally and pressing the correct key ($\bar{X}$ = 4.03) higher than either key pressing alone ($\bar{X}$ = 3.71) or vocalizing the number only ($\bar{X}$ = 2.85).

VOCAL STRESS ANALYSIS

29. The results of vocal stress analysis for the subgroup of 12 subjects are shown in figure 4. These data are mean plots of raw data uncorrected for individual baselines for each subject rank-ordered by percent-correct score (see paragraph 14). Figure 5 shows vocal stress corrected for each individual's baseline and grouped by presentation rate and rank order. The ANOVA of the vocal stress baseline deviation scores reveals significant main effects for rank order groups F = 3.97, df = 2/54, p < .05). Additional Tukey's mean comparison tests of these data show the middle-ranked group to be significantly different from the two other extreme-ranked groups at each presentation rate but only when required to press the correct key and vocalize the number. No other differences exist between rank-ordered groups under other response conditions.

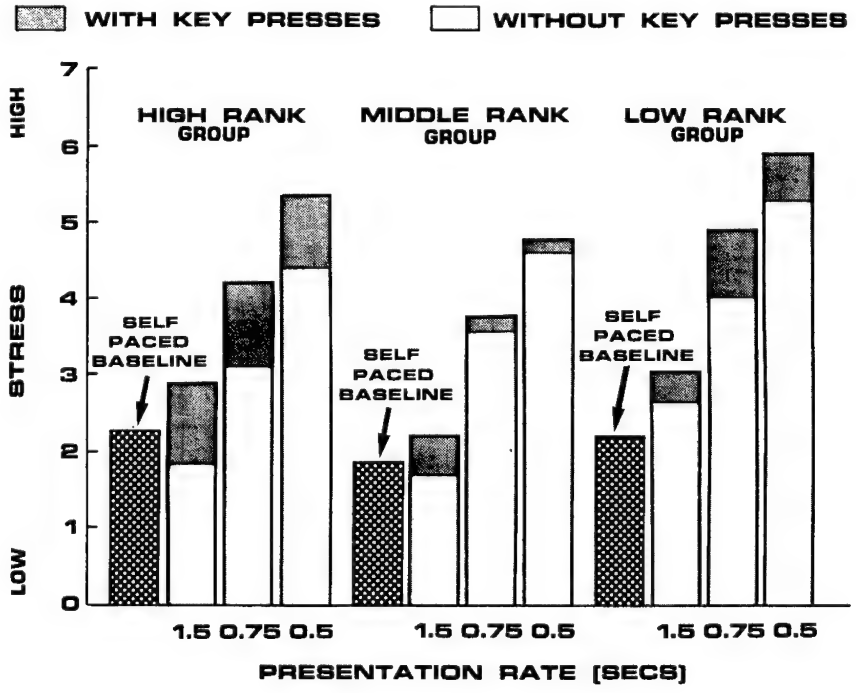


Figure 4
Measured Vocal Stress

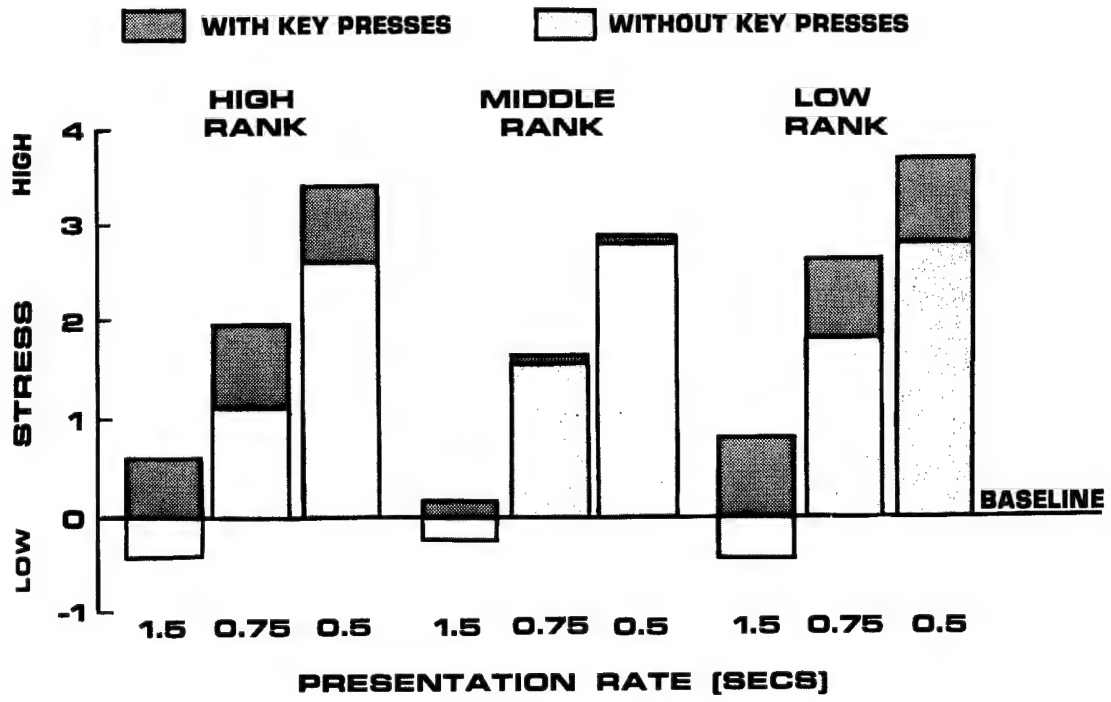


Figure 5
Measured Vocal Stress Deviations from Baseline

30. Presentation rate significantly affected deviations from baseline stress as measured in the voice across all response conditions ($F = 59.65$, $df = 2/54$, $p < .01$) except at the 1.5-sec presentation rate. Vocal stress significantly increased with simultaneous key pressing across rank-ordered groups and presentation rates $F = 8.96$, $df = 2/54$, $p < .01$) except with the middle-ranked group.

31. The pretest and posttest baseline conditions did not differ in vocal stress. Also, the baseline vocal stress measures taken during self-paced presentation rates showed no consistent trends with the automatic-paced presentation rates. That is, a subject that had a high or low baseline vocal stress level did not have a corresponding high or low vocal stress level during testing conditions. Therefore, the pretest baseline vocal stress measure could not be used as a predictor of the succeeding vocal stress levels during presentation rate increases.

CORRELATIONS

32. Spearman rank-order (Rho) correlations were calculated to determine if any rank-order relationships existed between subjects for percent-correct scores, self-rated stress, and vocal stress measures. Table I presents the resultant vocal stress rank-order correlations. Significant inverse relationships are shown to exist between vocal stress and percent-correct scores at the .75-sec and .5-sec rates, i.e., as percent-correct scores decreased, vocal stress increased. There were no significant correlations under "key pressing only" response conditions.

Table I

Rank-Order Correlations
(Vocal Stress X % Correct Scores)

VOCAL STRESS	% CORRECT SCORES					
	VOICE + KEY PRESSES			KEY PRESSES ONLY		
	1.5 SEC	0.75 SEC	0.5 SEC	1.5 SEC	0.75 SEC	0.5 SEC
VOICE ONLY	-0.13	-0.58*	-0.51*	0.03	-0.46	-0.27
VOICE + KEY PRESSES	0.06	-0.73**	-0.69*	-0.17	-0.48	-0.34

*RHO = 0.506, N = 12, $p < 0.05$ (ONE-TAILED TEST)
**RHO = 0.712, N = 12, $p < 0.01$ (ONE-TAILED TEST)

33. Significant rank-order correlations were found between vocal stress and self-rated stress as shown in table II at the .75-sec presentation rate but only when response conditions were identical, i.e., voice only or voice plus key presses. For example, vocal stress measurements taken during voice only response conditions showed no significant rank-order correlations with self-rated stress during voice plus key pressing response conditions.

Table II

Rank-Order Correlations
(Vocal Stress X Self-Rated Stress)

VOCAL STRESS	SELF-RATED STRESS					
	VOICE ONLY			VOICE + KEY PRESSES		
	1.5 SEC	0.75 SEC	0.5 SEC	1.5 SEC	0.75 SEC	0.5 SEC
VOICE ONLY	0.14	0.56*	0.23	0.02	0.42	0.02
VOICE + KEY PRESSES	0.04	0.44	0.19	0.16	0.60*	0.15

*RHO = 0.506, N = 12, $p < 0.05$ (ONE-TAILED TEST)

DISCUSSION

34. The psychomotor performance results of this study are consistent with previous findings by Kodalen (1975) and Helm, Fishburne, and Waag (1974) in using the RATER to deteriorate performance by increasing the presentation rate. That is, percent-correct scores decreased and self-rated stress increased as the requirement to process bits of information increased. However, the uniqueness of this study has shown that an objective, reliable, and sensitive measure can be utilized to study the underlying relationship between stress as manifested in the voice and performance on a psychomotor task.

35. The objectivity of this vocal measure has been established by identifying and quantifying the degree of correspondence that performance has with a subjective pattern recognition criterion of uniformity. In addition, this numerical rating technique was used to reliably indicate the degree of sensitivity present in the measure by scaling a range of uniformity over a wide spectrum of simulated-stress, waveform patterns.

36. These data indicate that the most predominate variable to affect all subjects under each response mode for each dependent measure was presentation rate. However, a closer examination of the data reveals that almost all of the dependent measures that reflected significant differences between groups did so only at the .75-sec presentation rate. This selective effect was due primarily to understimulation at the 1.5-sec rate and overstimulation at the .5-sec rate that produced an overload information processing situation. The percent-correct scores and self-rated stress estimates support the conclusion that the slowest and fastest rates narrowed the sensitivity range of these dependent measures to the extent that no discrimination could occur between the subjects, e.g., self-rated stress estimates were approximately the same. However, the vocal stress measure significantly differentiated between subjects at both the .75-sec and .50-sec rates which indicates a more sensitive measurement technique.

37. The sensitivity of the vocal measure to discriminate between the performance of subjects even when the percent-correct score range of 20.5% was narrow (38% maximum and 17.5% minimum) was further demonstrated by a highly significant correlation (Rho = -.69) when compared to the nonsignificant correlation of the self-rated stress estimates at the .50-sec rate. This lack of correlation is an inherent limitation imposed by any rating scale that is anchored at the extremes by either a maximum or minimum value. That is, the subjects were not allowed to manifest the absolute degree of stress. Likewise, it should be noted that vocal stress correlations with percent-correct scores (Rho = -.73) and self-rated stress estimates (Rho = .60) improved at the .75-sec rate. The corresponding percent-correct scores also increased from a range of 21% to a range of 52% and the self-rated estimates increased from a range of 0 to 4 stress units, indicating an improvement to discriminate or estimate individual performance. The degree of sensitivity of the vocal stress measure is an important finding of this study.

38. The correlational data also clearly indicate that performance can be better predicted from vocal stress measures or self-rated stress when the measures or ratings are taken during the same time of testing and involve the same reponse

mode, e.g., voice mode only. Vocalizing the numbers only resulted in significantly lower self-rated stress estimates and vocal stress measurements than when vocalizing and key pressing simultaneously. The data clearly reveal that key pressing was the main factor that increases stress. Whether the subjects would show a corresponding improvement or deterioration in vocal performance could not be evaluated because a reliable method of scoring correct verbal responses could not be devised due to the high rate of vocalization, e.g., .5 sec.

39. Another weakness in the study that limited a more extensive, statistical evaluation was the large number of data points per subject that necessitated scoring a smaller (N=12) sample than desired. The results and conclusions of this study should be interpreted with the reminder that the subjects were selected to represent the extremes and middle of the performance continuum for the purpose of proper scaling the sensitivity of a new measure. It is presently unknown whether the significant high correlations found in this study would remain significant with the inclusion of the remaining 30 subjects.

40. In summary, the following conclusions and recommendations are warranted based on the results of this study.

CONCLUSIONS

41. An objective, reliable, and sensitive pattern-recognition-scoring technique has been identified.

42. The quantification of this scoring technique has been automated.

43. The relationship of vocal stress measures to percent-correct scores and self-rated stress has been significantly correlated.

44. A potentially useful and unobtrusive test and evalaution technique has been established.

RECOMMENDATIONS

45. Plan collaborative efforts with other laboratories attempting to validate the Psychological Stress Evaluator.

46. Obtain field test data from aircrew tasks that require vocalizations over a range of information processing loads.

47. Scale the selected aircrew tasks by degree of vocalized stress.

48. Utilize the vocal stress measure in conjunction with secondary task methodology.

AUTOMATED VOCAL STRESS SCORING TECHNIQUE

GLOSSARY

Glossary of terms for the automatic stress scoring program flow diagram:

TRIGR - Trigger; the minimum value to be considered a signal. All values below this chosen level are considered to be noise and are disregarded by the computer.

ARRAY1 - Storage array for raw data.

ARRAY2 - Storage array for maximum and minimum values of the raw data from ARRAY1.

ARRAY3 - Storage array for the absolute value of the differences between consecutive values in ARRAY2, i.e., "peak-to-peak" values.

ST1 - Stress 1; a single number formed by taking the product of the differences between the values in the odd-numbered locations of ARRAY3. This value is related to the stress content contributed by either the positive or negative slopes of the waveform and is complimentary to ST2. Absolute determination of positive or negative slope is not relevant, just magnitude being significant.

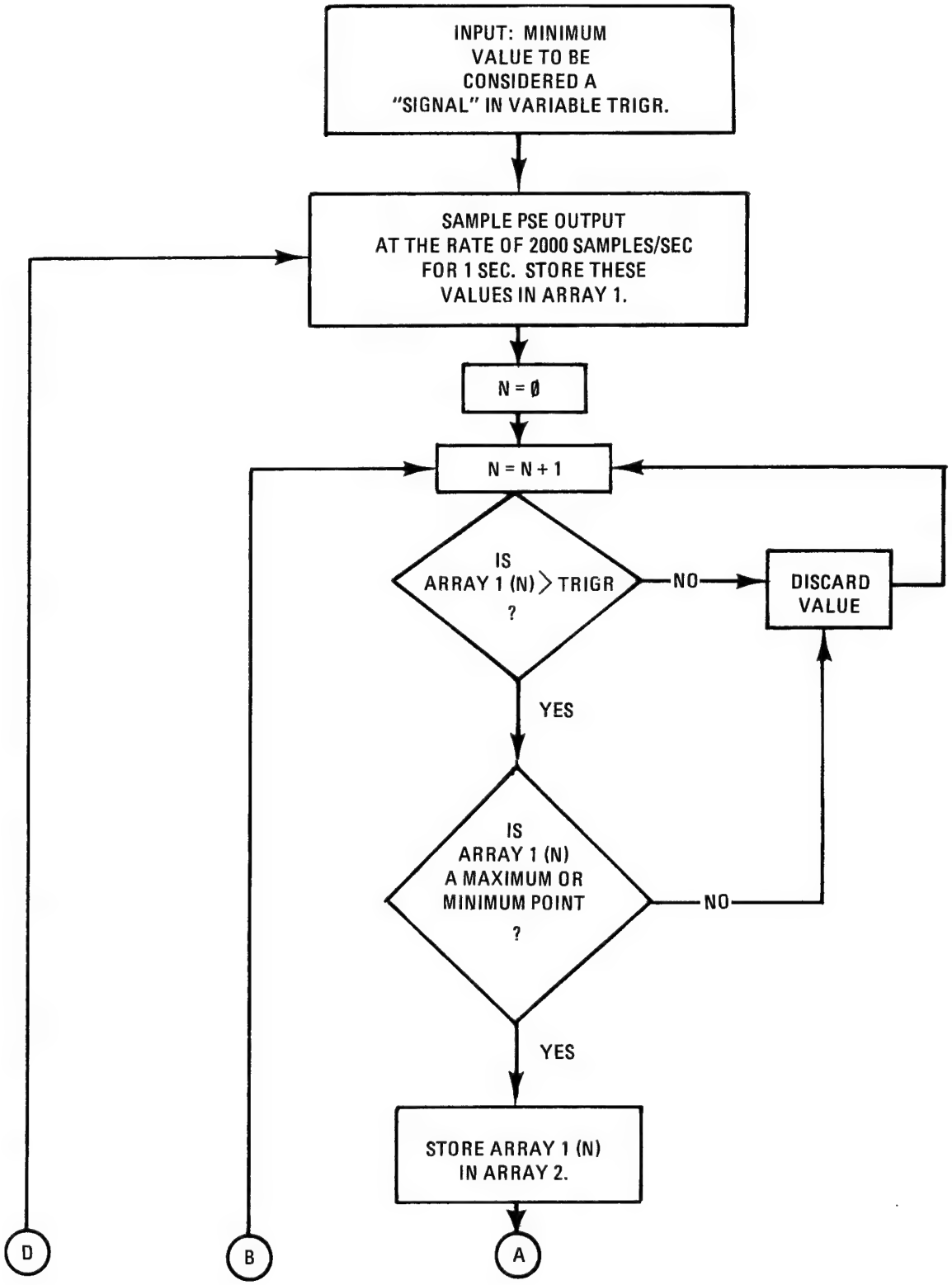
ST2 - Stress 2; the compliment of ST1.

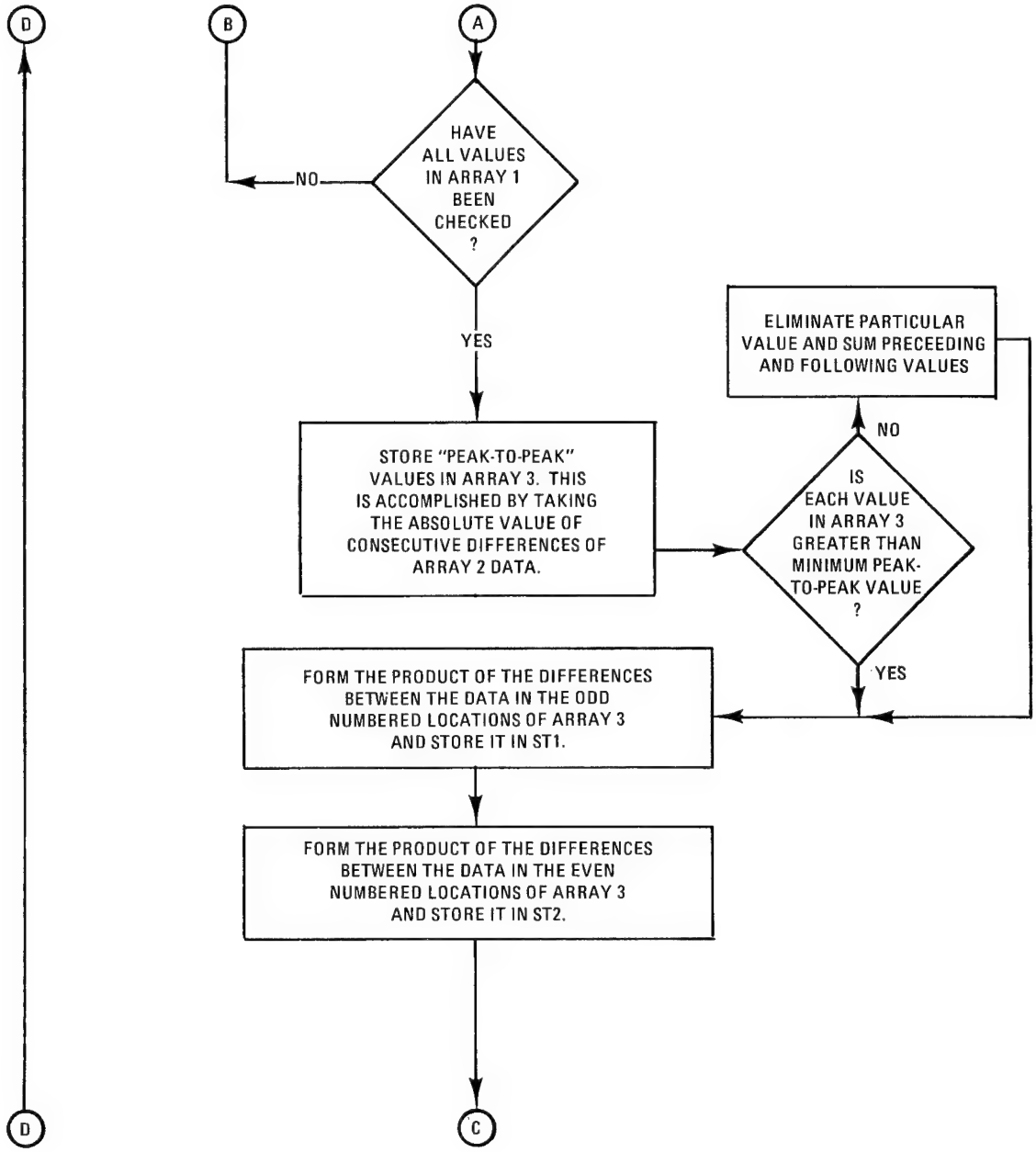
T1 - Tendency to Stress 1; a more workable form of ST1. $T1 = -\log_{10} ST1$. Larger values of T1 indicate higher stress.

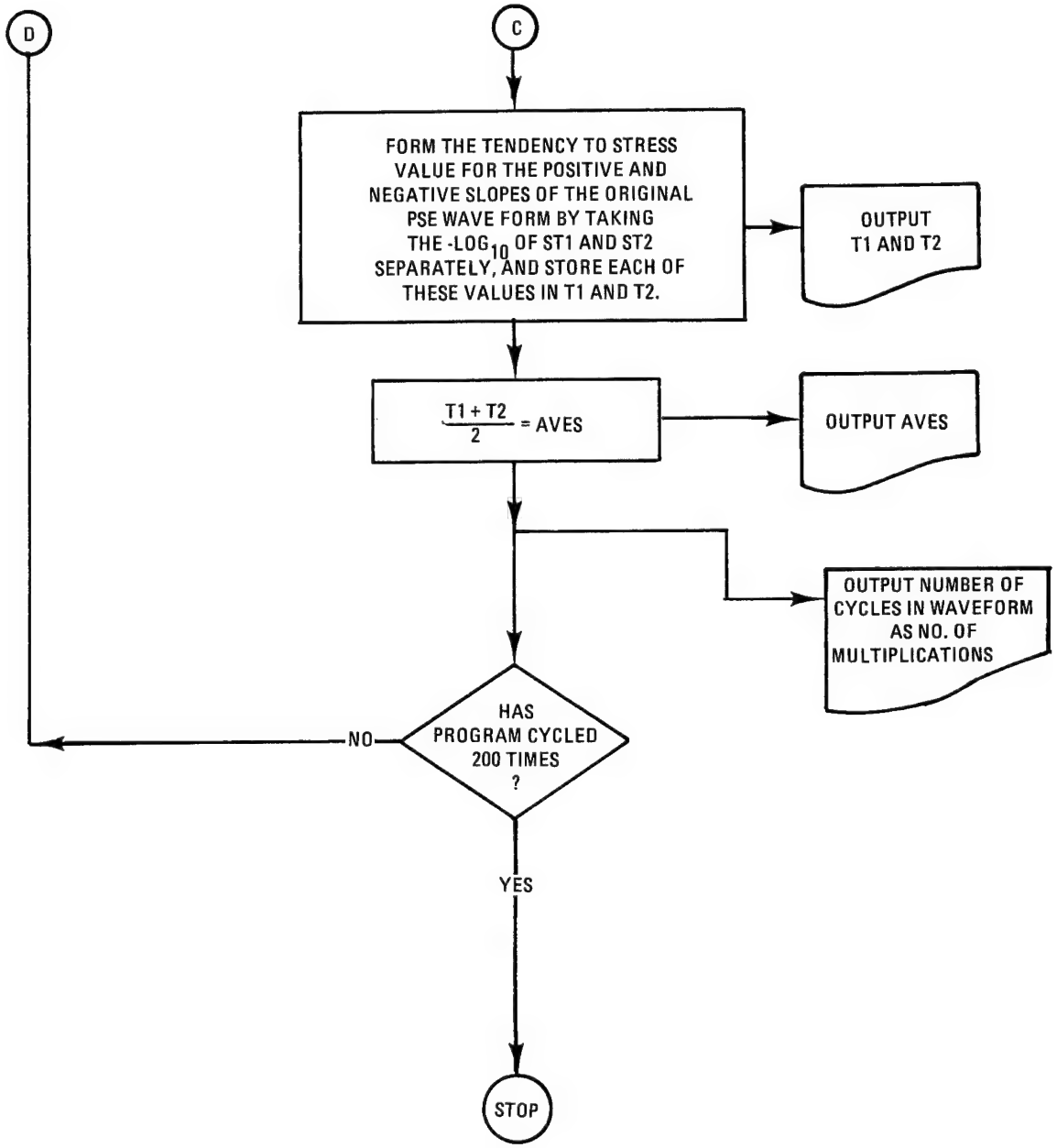
T2 - Tendency to Stress 2; a more workable form of ST2. $T2 = -\log_{10} ST2$. Larger values of T2 indicate higher stress.

AVES - Average Stress; the average of T1 and T2. One value of AVES should be calculated per each analyzed utterance.

AUTOMATIC STRESS SCORING PROGRAM FLOW DIAGRAM




D
B
A
HAVE ALL VALUES IN ARRAY 1 BEEN CHECKED ?
NO
YES
ELIMINATE PARTICULAR VALUE AND SUM PRECEEDING AND FOLLOWING VALUES
NO
STORE "PEAK-TO-PEAK" VALUES IN ARRAY 3. THIS IS ACCOMPLISHED BY TAKING THE ABSOLUTE VALUE OF CONSECUTIVE DIFFERENCES OF ARRAY 2 DATA.
IS EACH VALUE IN ARRAY 3 GREATER THAN MINIMUM PEAK-TO-PEAK VALUE ?
YES
FORM THE PRODUCT OF THE DIFFERENCES BETWEEN THE DATA IN THE ODD NUMBERED LOCATIONS OF ARRAY 3 AND STORE IT IN ST1.
FORM THE PRODUCT OF THE DIFFERENCES BETWEEN THE DATA IN THE EVEN NUMBERED LOCATIONS OF ARRAY 3 AND STORE IT IN ST2.
D
C


D
C
FORM THE TENDENCY TO STRESS VALUE FOR THE POSITIVE AND NEGATIVE SLOPES OF THE ORIGINAL PSE WAVE FORM BY TAKING THE -LOG10 OF ST1 AND ST2 SEPARATELY, AND STORE EACH OF THESE VALUES IN T1 AND T2.
OUTPUT T1 AND T2
T1 + T2 / 2 = AVES
OUTPUT AVES
OUTPUT NUMBER OF CYCLES IN WAVEFORM AS NO. OF MULTIPLICATIONS
HAS PROGRAM CYCLED 200 TIMES ?
NO
YES
STOP

REFERENCES

1. Bell, A. D., Jr., Ford, W. H., and McQuiston, C. R., Physiological Response Analysis Method and Apparatus, 1972, U.S. Patent Office (No. 3,971,034 issued 20 Jul 1976).

2. Borgen, L. A. and Goodman, L. I., Voice Print Analysis of Anxiolytic Drug Effects: Preliminary Results, Proceedings American Society of Clinical Pharmacology and Therapeutics, Seattle, Mar 1976, 19(104).

3. Branscomb, H. H., The Development of a Measure for Task-Induced Stress in Speech, Massachusetts Institute of Technology, Master of Science Thesis, 1979.

4. Brenner, M., Branscomb, H. H., and Schwartz, G. E., Psychological Stress Evaluator: Two Tests of a Vocal Measure, Psychophysiology, 1979, 16:351-357.

5. Brenner, M. and Branscomb, H. H., Psychological Stress Evaluator: Technical Limitations Affecting Lie Detection, Polygraph, 1979, 8:127-132.

6. Brenner, M., Branscomb, H. H., and Wright, R., A Comptuer Based Measure of Vocal Stress (Personnel Communication with NASA-Ames), 1979.

7. Cannings, R., Speech Patterns and Aircrew Workload, NATO AGARDOGRAPH #246, Aug 1979, pp 115-129.

8. Edson, R. K., The Dektor Psychological Stress Evaluator (Voice Stress Analyzer) as a Research Instrument, National Graduate University, 1976, Unpublished Master's Thesis, pp 1-113.

9. Gartner, W. B. and Murphy, M. R., Pilot Workload and Fatigue: A Critical Survey of Concepts and Assessment Techniques, NASA/Ames, NASA TN D-8365, Nov 1976.

10. Harrison, E. A., Stress Factors on Pilot Performance: A Bibliography with Abstracts, NTIS/PS-77/1160, Dec 1977, p 163.

11. Heisse, J. W., Jr., Is the Micro-Muscle Tremor Useable? The Micro-Muscle Tremor in the Voice, Hearing, 93rd Congress (2nd Session), June 1974, Wash: Government Printing Office, pp 339-342.

12. Heisse, J. W., Jr., Audio Stress Analysis of a Validation and Reliability Study of the Psychological Stress Evaluator (PSE), Carnahan Conference on Crime Countermeasures, May 1976, pp 5-17.

13. Helm, W. R., Fishburne, R. P., and Waag, W. L., The Effects of Verbal Interference on Psychomotor Performance under Dual Task Conditions, Proceedings Human Factors Society, 18th Annual Meeting, 1974, pp 471-475.

14. Hirsch, L. P. and Wiegele, T. C., Validity of the Psychological Stress Evaluator for Analyzing Voice Recordings in Political Research, Proceedings International Society of Political Psychology, May 1979, Washington, D.C.

15. Inbar, G. F. and Eden, G., Psychological Stress Evaluators: EMG Correlation with Voice Tremor, Biological Cybernetics, 1976, 24, p 165-7.

16. Kahneman, D., Tursky, B., Shapiro, D., and Crider, A., Pupilary, Heart Rate, and Skin Resistance Changes During a Mental Task, Journal of Experimental Psychology, 1969, 79, pp 164-167.

17. Kirk, R. E., Experimental Design: Procedures for the Behavioral Sciences, Brooks/Cole: California, 1968.

18. Kodalen, K. C., The Effect of Complex Information Processing Techniques Applied to Short Term Memory of Military Officers, ADA016376, Sep 1975.

19. Kubis, J. F., Comparison of Voice Analysis and Polygraph as Lie Detection Procedures, Hearing, 93rd Congress (2nd Session), Jun 1974, Wash: Government Printing Officer, pp 503-555.

20. Lieberman, P., Perturbations in Vocal Pitch, Journal of the Acoustical Society of America, 1961, 33: 597-603.

21. Lieberman, P. and Michaels, S. B., Some Aspects of Fundamental Frequency and Envelope Amplitude as Related to the Emotional Content of Speech, Journal of the Acoustical Society of America, 1962, 34: 922-927.

22. Lippold, O., Physiological Tremor, Scientific American, 1971, 224(3), pp 65-73.

23. McGlone, R., Test of the Psychological Stress Evaluator (PSE) as a Lie and Stress Detector. Paper presented at the Carnahan Conference on Crime Countermeasures, 1975.

24. Older, H. J. and Jenney, L. L., Psychological Stress Measurement Through Voice Output Analysis, NASA/JSC, N75-19960, 1975, p 56.

25. Planar Corporation, Results of Preliminary Tests of the PSE-1 Voice Analyzer, NASA/JSC, NAS 9-14146, 1972.

26. Roscoe, A. H. (Ed), Assessing Pilot Workload, AGARD-AG-233, Feb 1978.

27. Schiflett, S. G., Operator Workload: An Annotated Bibliography, Patuxent River, Maryland: U.S. Naval Air Test Center, SY-257R-76, Dec 1976.

28. Schiflett, S. G., Voice Stress Analysis as a Measure of Operator Workload, Proceedings of the Human Factors Society, 23rd Annual Meeting, Boston, Massachusetts, 29 Oct - 1 Nov 1979 (Abstract Only).

29. Siegel, S., Nonparametric Statistics, McGraw-Hill: New York, 1956.

30. Wierwille, W. W., Williges, R. C., and Schiflett, S. G., Aircrew Workload Assessment Techniques, NATO Advisory Group for Aerospace Research and Development, 1979, AGARDOGRAPH No. 246, pp 19-54.

DISTRIBUTION:

Wright-Patterson AFB (AFFDL/FGR)	(2)
Wright-Patterson AFB (AMRL/TSZ)	(1)
Wright-Patterson AFB (AMRL/HEB)	(3)
Army Research Institute, Fort Benning	(1)
NAVTRAEQUIPCEN	(1)
DARPA/CTO, Arlington	(1)
U.S. Army Research Institute of Environmental Medicine	(1)
U.S. Army AMRL, Fort Rucker	(1)
Brooks AFB (AFHRL/MPAP)	(1)
Brooks AFB (USAFSAM/VNE)	(3)
Brooks AFB (SAM/VNE)	(1)
NASA-AMES Research Center	(3)
PACMISTESTCEN (Code 1226)	(1)
Mental Health Center, Adelphia	(1)
Systemetrics, Incorporated, Blacksburg	(1)
Edwards AFB (AFFTC/DOEEH)	(1)
Naval Post Graduate School, Monterey	(1)
NAMRL, New Orleans	(1)
NASA, Langley Research Center	(1)
NAVAIRDEVCEN (Code 6021)	(1)
Office of Naval Research (Code 455)	(1)
U.S. Army HEL, Aberdeen Proving Ground	(1)
U.S. Army AMRL, Fort Rucker	(1)
White Sands Missile Range, New Mexico	(1)
HQ TCATA, Fort Hood	(1)
Office of Naval Personnel	(1)
NAVSEASYSCOMHQ	(1)
Naval Aerospace Medical Research Laboratory, Pensacola	(1)
Massachusetts Institute of Technology, Cambridge (Mr. M. E. Connelly)	(1)
Wright State University, Dayton (Dr. F. T. Eggemeier)	(1)
Hughes Helicopter, Culver City (Mr. S. A. Thompson)	(1)
Rockwell International/Collins Division, Cedar Rapids (Dr. L. Silverstein)	(1)
Dr. Jack Laveson, Annandale, Va.	(1)
Northern Illinois University, Dekab (Dr. T. Wiegele)	(1)
University of South Dakota, Vermillion (LCDR W. Helm)	(1)
Dektor, Incorporated, Springfield (Ms. P. Gerner)	(3)

Systems Research Laboratories, Incorporated, Dayton (Mr. R. Spicuzza)	(1)
Bell Helicopter Company, Fort Worth (Dr. D. Strothers)	(1)
University of Illinois, Champaign (Dr. C. Wickens)	(1)
Honeywell, Incorporated, Minneapolis (Dr. R. North)	(1)
Systems Research Laboratories, Incorporated, Dayton (Mr. M. Crabtree)	(1)